

Federal Trade Commission Public Workshop on Consumer Information Privacy

*Supplemental Comments for
Project Number: P954807
Document Number: 18*

James E. Pitkow
Xerox Palo Alto Research Center
Palo Alto California 94304
phone: 415-812-4941
fax: 415-812-4241

Colleen M. Kehoe
Georgia Institute of Technology
Atlanta Georgia 30332
phone: 404-894-6266
fax: 404-894-0673

Overview of Comments

GVU's WWW Data Privacy User Surveys

The comments contained in this document are being submitted as a supplement to Project Number P954807 Document Number 18 by the same authors. These comments are intended to document recent developments resulting from GVU's Seventh World Wide Web User Survey conducted April 10 through May 10, 1997. The survey received over 19,000 respondents. The new results are relevant to the following areas being discussed in Session Two on Consumer Online Privacy:

1. Web sites' current actual practices in the collection, compilation, sale, and use of consumers' personal information;
2. current design and implementation of technologies intended to enhance online information privacy; and
3. unsolicited commercial e-mail.

The results presented are preliminary. Subsequent analysis is being performed and will be made publicly available from the GVU User Survey Web Site

http://www.gvu.gatech.edu/user_surveys

starting June 10, 1997 along with presented during Session Two of the Workshop. More sophisticated analyses will be conducted over the summer as well. As always, we welcome comments on the results and suggested refinements to the set of questions asked.

Session Two: Consumer Online Privacy

Information Collection and Use

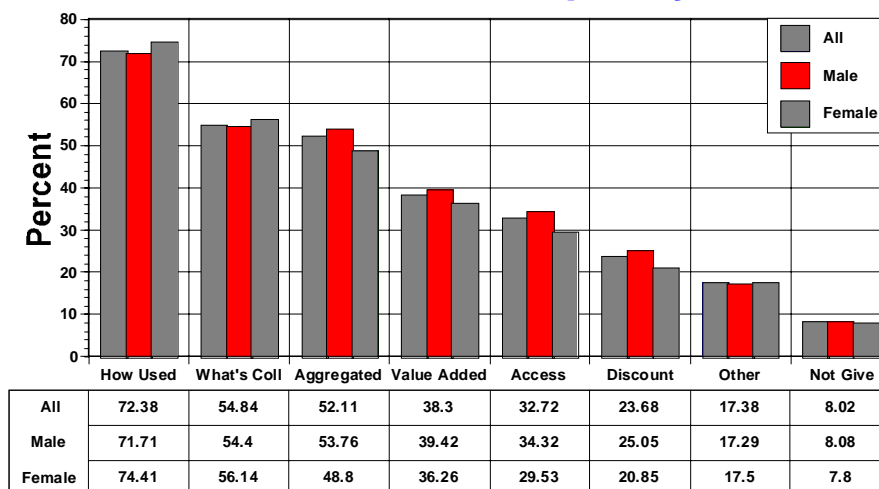
Question: 2.2

To what extent is the collection, compilation, sale or use of personally identifying, as opposed to aggregate, personal information important for marketing online and for market research? What privacy concerns, if any, are raised by the collection or use of aggregate personal information in this context?

Comment:

Results from the most recent survey are consistent with the results previously reported. Just over 52% of respondents replied that they are willing to provide personal information to web sites when the information was only used in aggregate. While this is a slight decrease from the previous survey where 55% were willing to provide this kind of information, it is seen across both genders and all age groups. The largest change is in the elder generation from 64% in the 6th survey to 56% in the 7th. In the previous survey, we noted that the elder generation showed a stronger preference for data to be used in aggregate than the younger generation (an 11% difference), but this difference has decreased to 7% in the most recent survey.

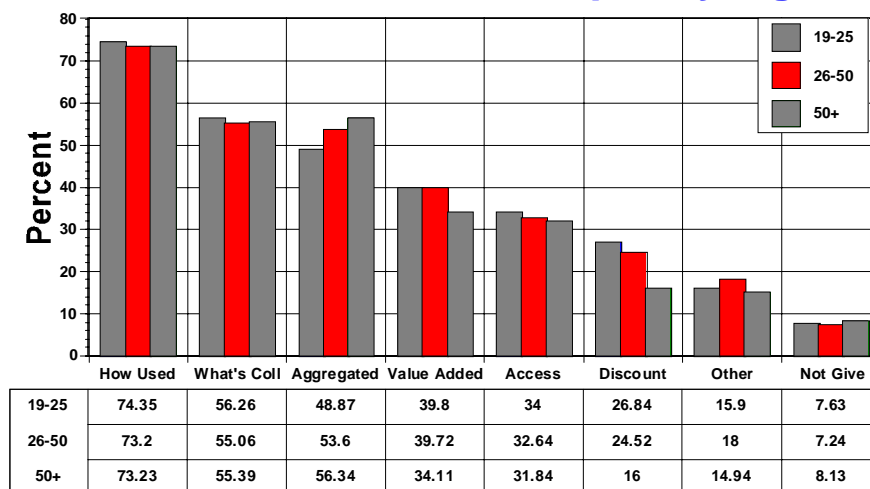
Terms and Conditions split by Gender



Source: Gvu's Seventh WWW User Survey™ (Conducted April 1997)
 <URL: http://www.gvu.gatech.edu/user_surveys/>
 Copyright 1997 GTRC - ALL RIGHTS RESERVED
 Contact: www-survey@cc.gatech.edu

Figure 1. The terms and conditions for revealing personal information to WWW sites. Respondents would provide sites with information under the following conditions: a statement were provided as to how the information was going to be used ("How Used"), a statements was provided about what information was being collected ("Info Gotten"), if the collected data were only used in aggregate ("Aggregated"), in exchange for some value added service ("For Service"), in exchange for access to the site ("Access to Site"), for a discount at the site's store ("Discount"), for some other reason ("Other"), and that they would not provide personal information to a site ("Not Do"). Respondents were allowed to check more than one answer for this question.

Terms and Conditions split by Age



Source: Gvu's Seventh WWW User Survey™ (Conducted April 1997)
 <URL: http://www.gvu.gatech.edu/user_surveys/>
 Copyright 1997 GTRC - ALL RIGHTS RESERVED
 Contact: www-survey@cc.gatech.edu

Figure 2. As with the Sixth Survey, there is basic agreement across age segment about the terms and conditions necessary to reveal demographic information. Respondents 50 yr. and older show a stronger preference for the collected information to be used in aggregate, a finding that was also found in the Sixth Survey.

Session Two: Consumer Online Privacy

Information Collection and Use

Question: 2.4

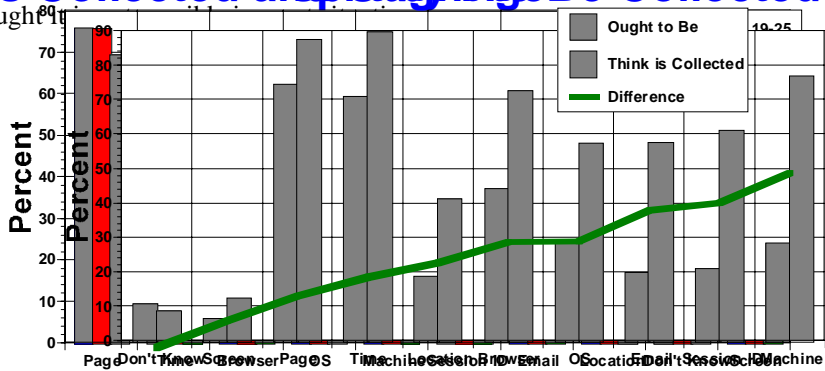
What surveys, other research, or quantitative or empirical data exist about consumers' perceptions, knowledge and expectations regarding (1) whether their personal information is being or should be collected by Web site operators and the extent of such collection; (2) the benefits and risks associated with the collection and subsequent use of this information; (3) appropriate uses of such information; and (4) whether certain categories of information should never be collected or disclosed to others?

Comment:

When asked what information should be allowed to be collected over the WWW, respondents in the most recent survey gave answers that were virtually identical to those from the last survey. The majority of respondents are comfortable with sites recording what page is requested and the time of the request (over 70% agree). Respondents are somewhat less comfortable with the kind of browser (44%), machine name (28%) and operating system (28%) being collected and are considerably less comfortable with more personal information being collected with email address, location, and session ID having 20% or less agreement. The choice of "screen size" was added to this question for the most recent survey since it can be collected through some web technologies (e.g. Java). Only 6% of respondents felt it should be collected.

In a longitudinal analysis of 160 respondents who participated in both the 6th and 7th surveys, we notice that after 6 months, these respondents have become somewhat better educated about what information might be collected during a WWW transaction. The percentage of users who recognized that session IDs, their kind of browser and kind of operating system might be collected each increased by 5%. Additionally, the percentage of users who thought physical location might be recorded dropped by 7%. Physical location was the only choice provided which clearly cannot be gathered reliably on the WWW. Confusion persists, however, about whether email addresses can be collected. The percentage of users who thought they could be collected increased from 19% to 61% even though

Difference Between What Collected and What Users Think is Collected and What Users Think Should Be Collected



Opinions on Direct Marketing split by Location



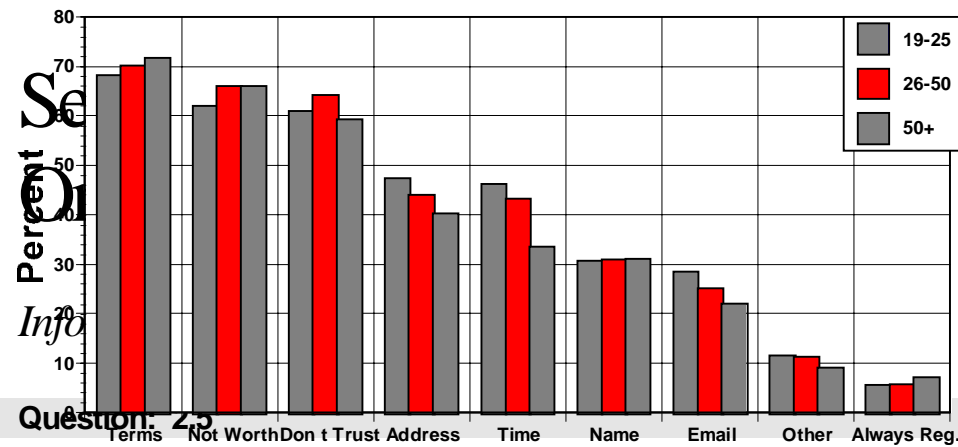
Figure 4. People's responses to what information they think ought to be collected and what is currently collected by age. The categories are as follows: the page that is being requested ("Page"), the time the page is requested ("Time"), the type of browser being used ("Browser"), the domain name of the machine making the request ("Machine"), the type of operating system the user is using ("OS"), the user's name ("Name"), the user's email address ("Email"), the location of the user making the request ("Location"), an identifier that persists across sessions ("Session ID") and finally those users that did not know what information ought to be collected ("Don't Know").

Source: GVV's Seventh WWW User Survey (Conducted April 1997)
 URL: http://www.gvu.gatech.edu/user_surveys/
 Copyright 1997 GTRC - ALL RIGHTS RESERVED
 Contact: www-survey@cc.gatech.edu

Figure 5. Respondents were asked to state their agreement with certain statements about personal information collection and direct marketing. Although the graphic is stratified by gender, there is little disagreement between strata for each statement.

Figure 3. Clear differences exist between what people would like to be collected and what they think is collected. Overall, users are not very well educated about that information is currently being collected, as all the above data points except a user's email address are collectable by common practice.

Reason for Not Registering split by Age



Reason	19-25	26-50	50+
Terms	68.15	70	71.65
Not Worth	61.91	65.99	65.95
Don't Trust	60.99	64.07	59.29
Address	47.24	44.03	40.18
Time	46.04	43.15	33.42
Name	30.62	30.88	31.05
Email	28.39	25.15	21.96
Other	11.41	11.23	9.08
Always Reg.	5.57	5.82	7.02

Source: GVVU's Seventh WWW User Survey (Conducted April 1997)

Copyright 1997 GTRC - ALL RIGHTS RESERVED

Contact: www-survey@cc.gatech.edu

Comment:

Figure 6. When asked why users do not register at sites that request personal information, the number one reason is the lack of clearly specified terms and conditions ("Terms"), followed by a general perception that it is not worth it ("Too Risky"), that users do not trust the Web site requesting the information ("Don't Trust"), that the site is requesting address information ("Address"), the amount of time it takes to provide the information ("Time"), that the site requesting the user's name ("Name"), electronic mail address ("Email"), some other reason ("Other"), and a group of users who ("Always") register when asked by a site.

such as their name, postal address, and email address is troubling for 30% to 50% of respondents.

Session Two: Consumer Online Privacy

Unsolicited Commercial Email

Question: 2.17

What are the risks and benefits, to both consumers and commercial entities, of unsolicited commercial e-mail? What are consumers' perceptions, knowledge, and expectations regarding the risks and benefits of unsolicited commercial e-mail?

Comment:

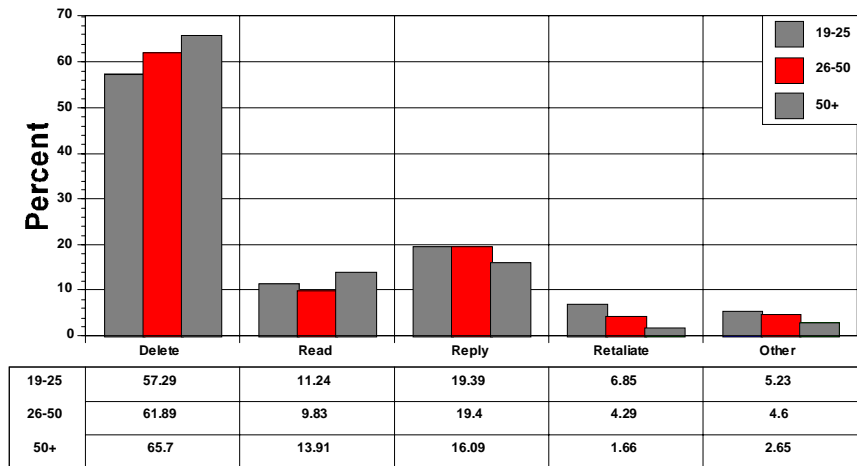
A longitudinal analysis of participants in both surveys reveals an interesting difference after six months. The percentage of respondents who disagreed strongly that they liked receiving mass emailings increased from 63% in the 6th survey to 74% in the 7th survey. There is a corresponding, even larger increase in the percentage that disagreed strongly that they liked receiving mass postal mailings: 30% in the 6th survey to 48% in the 7th survey. While this is only one data point, it suggests that users may become less tolerant of mass emailings over time. It is unclear how this is related to a decreased tolerance for mass postal mailings.

The percentage of respondents who simply delete the message rose slightly from 59% in the 6th survey to 61% in the 7th. The percentage that read the message showed an equivalent decrease from 13% in the 6th to 11% in the 7th. (The percentages given here for the 6th survey are different from those reported in the previous submission because respondents who reported that they had never received mass email were excluded from this most recent analysis.) Women and respondents over 50 continue to be more likely to simply delete mass email than other demographic groups.

Our longitudinal analysis of respondents who completed both the 6th and 7th surveys shows that they have become more likely to simply delete mass email when they receive it: 41% in the 6th survey and 58% in the 7th. Most of this change can be attributed to the fact that 14% reported in the 6th survey that they had never received mass emailings compared to only 6% in the 7th. The percentage of respondents who read the messages also decreased from 7% to 2%.

For the question of what to do about mass emailings, we added a new choice to the list for the most recent survey which was to create a "blacklist" of known spammers to allow email from them to be filtered out automatically. This choice ranked second with 19% of users preferring this option. Since the question changed between the two surveys, percentages are not directly comparable. It is worth noting, however, that the ordering of the top three options is the same for both surveys with an opt-out registry being the most preferred in both cases. An interesting change in the ordering is that having government regulation making spamming illegal out ranked "doing nothing" in the most recent survey. The difference between these two is still relatively small (2%), but it will be an important point to watch as Internet users search for ways to deal with mass emailings.

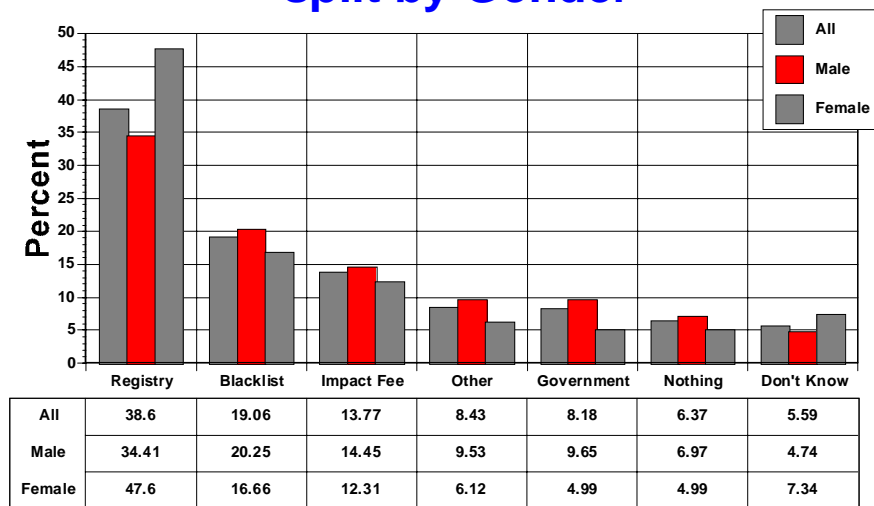
Attitude Toward Spamming split by Age



Source: GVU's Seventh WWW User Survey[™] (Conducted April 1997)
 <URL: http://www.gvu.gatech.edu/user_surveys/>
 Copyright 1997 GTRC - ALL RIGHTS RESERVED
 Contact: www-survey@cc.gatech.edu

Figure 7. People react to receiving mass emailings, a.k.a. spam, in different ways. The most common reaction is to "Delete" the message, followed by requesting to be removed from the list "Ask to be Removed", and some form of retaliation, "Retaliate", which is typically in the form of mail bombs, denial of service attacks, etc. Around 20% claim never to have received mass emailings "Not Applicable" and 10% or so claim to read the contents of the message "Read Message." While the older generation is more likely to read and delete messages, the younger generation is more likely to retaliate in some manner.

Policy Toward Spamming split by Gender



Source: GVU's Seventh WWW User Survey[™] (Conducted April 1997)
 <URL: http://www.gvu.gatech.edu/user_surveys/>
 Copyright 1997 GTRC - ALL RIGHTS RESERVED
 Contact: www-survey@cc.gatech.edu

Figure 8. Most users, especially female users, recommend the creation of a registry of users that do not wish to receive mass emailings. This registry would need to be consulted by entities wishing to broadcast messages. The other options included: imposing an "Impact Fee" on the origin entities, doing "Nothing" and some form of "Government Control."